

CLAIMS

We claim:

1. A method comprising at least one of:

assigning interrupts for a plurality of input/output (I/O) devices among a plurality of nodes
5 of a non-uniform memory access (NUMA) system based on at least one of: the nodes to which
the I/O devices are connected; the nodes at which interrupt service routines for the I/O devices
reside; and, processors of the nodes;

for each node of the NUMA system, assigning the interrupts for the devices that are
performance critical and that have been assigned to the node among the processors of the node
10 in a round-robin manner;

dynamically modifying assignments of the interrupts among the nodes of the NUMA system
based on actual performance characteristics of the assignments; and,

for each node of the NUMA system, dynamically modifying assignments of the interrupts
that are performance critical and that have been assigned to the node among the processors of
15 the node based on actual performance characteristics of the assignments.

2. The method of claim 1, wherein assigning the interrupts for the plurality of I/O devices
among the plurality of nodes of the NUMA system comprises, for each I/O device:

where the node to which the I/O device is connected has a cache, memory, and at least one
processor, assigning the interrupt for the I/O device to the node to which the I/O device is
20 connected; and,

otherwise, where the node at which the interrupt service routine for the I/O device resides
has memory and at least one processor, assigning the interrupt for the I/O device to the node at
which the interrupt service routine for the I/O device resides.

3. The method of claim 2, wherein assigning the interrupts for the plurality of I/O devices

among the plurality of nodes of the NUMA system further comprises, for each I/O device,
25 otherwise, assigning the interrupt for the I/O device to one of the nodes having memory and at
least one processor.

4. The method of claim 1, wherein dynamically modifying the assignments of the interrupts
among the nodes of the NUMA system comprises, for each assignment of an interrupt for an I/O

device to a node, where the node is that to which the I/O device is connected:

measuring responsiveness of the node in processing the interrupt;

assigning the interrupt to the node at which the interrupt service routine for the I/O device resides;

5 measuring responsiveness of the node at which the interrupt service routine for the I/O device resides in processing the interrupt; and,

where the responsiveness of the node to which the I/O device is connected is better than the responsiveness of the node at which the interrupt service routine for the I/O device resides, reassigning the interrupt to the node to which the I/O device is connected.

10 5. The method of claim 4, wherein dynamically modifying the assignments of the interrupts among the nodes of the NUMA system comprises, for each assignment of an interrupt for an I/O device to a node, where the node is that at which the interrupt service routine for the I/O device resides:

measuring responsiveness of the node in processing the interrupt;

15 assigning the interrupt to the node to which the I/O device is connected;

measuring responsiveness of the node to which the I/O device is connected in processing the interrupt; and,

where the responsiveness of the node at which the interrupt service routine for the I/O device resides is better than the responsiveness of the node to which the I/O device is connected,

20 reassigning the interrupt to the node at which the interrupt service routine for the I/O device resides.

6. The method of claim 1, wherein, for each node of the NUMA system, dynamically modifying the assignments of the interrupts that are performance critical and that have been assigned to the node among the processors of the node comprises:

25 measuring responsiveness of the processors of the node in processing the interrupts assigned thereto;

where a differential between a best responsiveness and a worst responsiveness is greater than a threshold,

30 reassigning at least one of the interrupts assigned to the processor having the worst responsiveness to the processor having the best responsiveness.

7. A non-uniform memory access (NUMA) system comprising:

a plurality of nodes;

a plurality of input/output (I/O) devices, each I/O device connected to one of the plurality of nodes and having an interrupt; and,

5 an interrupt-assignor responsive to the I/O devices and the nodes to assign the interrupt for each I/O device to one of the plurality of nodes in a performance-optimized manner.

8. The system of claim 7, wherein each of one or more of the plurality of nodes has memory and at least one processor, the memory of a node being local to the node and remote to other of the plurality of nodes, and the interrupt-assignor is to assign the interrupt for each I/O device to
10 one of the plurality of nodes that has memory and at least one processor.

9. The system of claim 8, wherein at least one of the I/O devices are performance critical, the interrupt-assignor further to assign the interrupt for each I/O device that is performance critical among the at least one processor of the node to which the interrupt has been assigned in a round-robin manner.

15 10. The system of claim 7, wherein the interrupt-assignment software is further to dynamically modify assignments of the interrupts that are performance critical among the at least one processor of a node based on actual performance characteristics of the assignments.

11. The system of claim 8, wherein each of one or more of the plurality of nodes is memoryless and processorless.

20 12. The system of claim 7, wherein the interrupt-assignor is further to dynamically modify assignments of the interrupts among the plurality of nodes based on actual performance characteristics of the assignments.

13. The system of claim 7, wherein the interrupt-assignor is to give primary preference in assigning the interrupt for each I/O device to the node to which the I/O device is connected
25 where the node to which the I/O device is connected has a cache, memory, and at least one processor.

14. The system of claim 13, wherein each I/O device further has an interrupt service routine residing at one of the plurality of nodes, and the interrupt-assignor is to give secondary preference in assigning the interrupt for each I/O device to the node at which the interrupt service routine of the I/O device resides where the node at which the interrupt service routine of the I/O device resides has a cache, memory, and at least one processor.

15. The system of claim 7, wherein the interrupt-assignor resides within one of the plurality of nodes.

16. An article of manufacture comprising:

a computer-readable medium; and,

means in the medium for assigning interrupts for a plurality of input/output (I/O) devices among a plurality of nodes based on at least one factor selected from the set consisting of: the nodes to which the I/O devices are connected; and the nodes at which interrupt service routines for the I/O devices reside.

17. The article of claim 16, wherein the means is for assigning the interrupts among the plurality of nodes further based on whether the nodes have processors and memories.

18. The article of claim 16, wherein the means, for each node, is further for assigning the interrupts for the devices that are performance critical and that have been assigned to the node among the processors of the node in a round-robin manner.

19. The article of claim 18, wherein the means, is further for dynamically modifying assignments of the interrupts among the nodes based on actual performance characteristics of the assignments, and, for each node, for dynamically modifying assignments of the interrupts that are performance critical and that have been assigned to the node among the processors of the node based on actual performance characteristics of the assignments.

20. The article of claim 16, wherein the medium is one of a recordable data storage medium and a modulated carrier signal.

21. An article of manufacture comprising

an interrupt-assignor to assign interrupts for a plurality of input/output (I/O) devices among a plurality of nodes based on at least one factor selected from the set consisting of: the nodes to which the I/O devices are connected; and the nodes at which interrupt service routines for the I/O devices reside.

22. The article of claim 21, wherein the interrupt-assignor is to assign the interrupts among the plurality of nodes further based on whether the nodes have processors and memories.

23. The article of claim 21, wherein the interrupt-assignor is to assign, for each node, the interrupts for the devices that are performance critical and that have been assigned to the node among the processors of the node in a round-robin manner.

24. The article of claim 23, wherein the interrupt-assignor is to dynamically modify assignments of the interrupts among the nodes based on actual performance characteristics of the assignments, and, for each node, to dynamically modify assignments of the interrupts that are performance critical and that have been assigned to the node among the processors of the node based on actual performance characteristics of the assignments.

25. A method comprising:

assigning interrupts for a plurality of input/output (I/O) devices among a plurality of nodes based on at least one factor selected from the set consisting of: the nodes to which the I/O devices are connected; and the nodes at which interrupt service routines for the I/O devices reside;

for each node of the system, assigning the interrupts for the devices that are performance critical and that have been assigned to the node among the processors of the node in a round-robin manner;

dynamically modifying assignments of the interrupts among the nodes of the system based on actual performance characteristics of the assignments; and,

for each node of the system, dynamically modifying assignments of the interrupts that are performance critical and that have been assigned to the node among the processors of the node based on actual performance characteristics of the assignments.

26. The method of claim 25, wherein assigning the interrupts for the plurality of I/O devices among the plurality of nodes of the system comprises, for each I/O device:

where the node to which the I/O device is connected has a cache, memory, and at least one processor, assigning the interrupt for the I/O device to the node to which the I/O device is

5 connected; and,

otherwise, where the node at which the interrupt service routine for the I/O device resides has memory and at least one processor, assigning the interrupt for the I/O device to the node at which the interrupt service routine for the I/O device resides.

27. The method of claim 26, wherein assigning the interrupts for the plurality of I/O devices among the plurality of nodes of the system further comprises, for each I/O device, otherwise, assigning the interrupt for the I/O device to one of the nodes having memory and at least one processor.

28. The method of claim 25, wherein dynamically modifying the assignments of the interrupts among the nodes of the system comprises, for each assignment of an interrupt for an I/O device to a node, where the node is that to which the I/O device is connected:

measuring responsiveness of the node in processing the interrupt;
assigning the interrupt to the node at which the interrupt service routine for the I/O device resides;

measuring responsiveness of the node at which the interrupt service routine for the I/O device resides in processing the interrupt; and,

where the responsiveness of the node to which the I/O device is connected is better than the responsiveness of the node at which the interrupt service routine for the I/O device resides, reassigning the interrupt to the node to which the I/O device is connected.

29. The method of claim 28, wherein dynamically modifying the assignments of the interrupts among the nodes of the system comprises, for each assignment of an interrupt for an I/O device to a node, where the node is that at which the interrupt service routine for the I/O device resides:

measuring responsiveness of the node in processing the interrupt;
assigning the interrupt to the node to which the I/O device is connected;
measuring responsiveness of the node to which the I/O device is connected in processing the interrupt; and,

where the responsiveness of the node at which the interrupt service routine for the I/O device

resides is better than the responsiveness of the node to which the I/O device is connected, reassigning the interrupt to the node at which the interrupt service routine for the I/O device resides.

30. The method of claim 25, wherein, for each node of the system, dynamically modifying the assignments of the interrupts that are performance critical and that have been assigned to the node among the processors of the node comprises:

measuring responsiveness of the processors of the node in processing the interrupts assigned thereto;

where a differential between a best responsiveness and a worst responsiveness is greater than a threshold,

reassigning at least one of the interrupts assigned to the processor having the worst responsiveness to the processor having the best responsiveness.